

ABSTRACT OF THE DISCLOSURE

In evaluating commonality of documents, each sentence is represented by a binary vector whose components indicate the presence or absence of corresponding terms, whereupon the concept of a common vector among documents is introduced. One sentence vector is derived from each of the documents to form a group of sentence groups, and only components which assume "1" (one) in all the vectors are "1", the other components being "0" (zero). The commonality of a document set is evaluated by employing the sum or squared sum of the numbers of components whose values are not zero in the individual common vectors, for all the common vectors.